The Journal of Veterinary Medical Science

FULL PAPER

Biochemistry

Dog Steroidogenic Factor-1: Molecular cloning and analysis of epigenetic regulation

Asato SEKIYA[1)], Ken TAKASAWA[1,4)], Yoshikazu ARAI[1)], Shidow TORISU[2)] and Koichiro NISHINO[1,3)]*

[1)]Laboratory of Veterinary Biochemistry and Molecular Biology, Graduate School of Medicine and Veterinary Medicine/Faculty of Agriculture, University of Miyazaki, 1-1 Gakuen-Kibanadai-Nishi, Miyazaki 889-2192, Japan

[2)]Veterinary Teaching Hospital, Faculty of Agriculture, University of Miyazaki, 1-1 Gakuen-Kibanadai-Nishi, Miyazaki 889-2192, Japan

[3)]Center for Animal Disease Control, University of Miyazaki, 1-1 Gakuen-Kibanadai-Nishi, Miyazaki 889-2192, Japan

[4)]Present address: Division of Molecular Modification and Cancer Biology, National Cancer Center, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan

ABSTRACT. Steroidogenic factor 1 (SF-1) is a nuclear receptor that is important in steroid hormone production, and adrenal and gonad development. The *SF-1* gene is highly conserved among most vertebrates. However, dog *SF-1* registered in public databases, such as CanFam3.1, lacks the 5′ end compared to other mammals including mouse, human, bovine, and cat. Whether this defect is due to species differences or database error is unclear. Here, we determined the full-length dog *SF-1* cDNA sequence and identified the missing 5′ end sequence in the databases. The coding region of the dog *SF-1* gene has 1,386 base pairs, and the protein has 461 amino acid residues. Sequence alignment analysis among vertebrates revealed that the 5′ end sequence of dog *SF-1* cDNA is highly conserved compared to other vertebrates. The genomic position of the first exon was determined, and its promoter region sequence was analyzed. The DNA methylation state at the basal promoter and the expression of dog *SF-1* in steroidogenic tissues and non-steroidogenic cells were examined. CpG sites at the basal promoter displayed methylation kinetics inversely correlated with gene expression. The promoter was hypomethylated and hypermethylated in *SF-1* expressing and non-*SF-1* expressing tissues, respectively. In conclusion, we identified the true full sequence of dog *SF-1* cDNA and determined the genome sequence around the first exon. The gene is under the control of epigenetic regulation, such as DNA methylation, at the promoter.

KEY WORDS: DNA methylation, dog, molecular cloning, promoter, Steroidogenic factor 1

J. Vet. Med. Sci.
82(6): 681–689, 2020
doi: 10.1292/jvms.20-0050

Received: 30 January 2020
Accepted: 24 March 2020
Advanced Epub:
2 April 2020

Steroid hormones are bioactive substances synthesized in the adrenal cortex and gonads of vertebrates. These hormones are indispensable for homeostasis, metabolism, and sexual differentiation. Representative steroid hormones include glucocorticoid and mineralocorticoid produced in the adrenal cortex, testosterone produced in the testis, and estradiol and progesterone produced in the ovary. The orphan nuclear receptor steroidogenic factor 1 (SF-1, also called Ad4BP or NR5A1) is expressed mainly in steroidogenic tissues [24] and regulates transcription of genes associated with steroidogenesis by binding to DNA sequences commonly preserved in the promoter regions [3–5, 10, 16, 18, 28–30, 34]. The observations that *SF-1* knockout mice lack adrenal glands and gonads and die soon after birth indicate that SF-1 plays an important role in fetal development and sex differentiation [15].

SF-1 contains several domains. An N-terminal DNA binding domain (DBD) includes two Cys2 -Cys2 zinc fingers (ZF-I and ZF-II) and Fushi-tarazu factor 1 (FTZ-F1); a flexible hinge region involved in protein structural variability; a C-terminal ligand binding domain (LBD); and two activation function domains (AF-1 and AF-2) [9]. The amino acid sequence of SF-1 is highly conserved among most vertebrates including human, mouse, rat, bovine, pig, and cat. However, the dog SF-1 sequence recorded in CanFam3.1 (ENSCAFT0000032206) lacks the sequence equivalents of exon 1 and 2 of humans and mice (Fig. 1). Compared with the genomic sequence at the basal promoter regions, several transcription factor binding sites are conserved widely in mammals, but not in dog (based on ENSCAFT00000032206) [6, 20, 27, 31]. Whether the deficient structure of SF-1 specific to dog is species

*Correspondence to: Nishino, K.: aknishino@cc.miyazaki-u.ac.jp

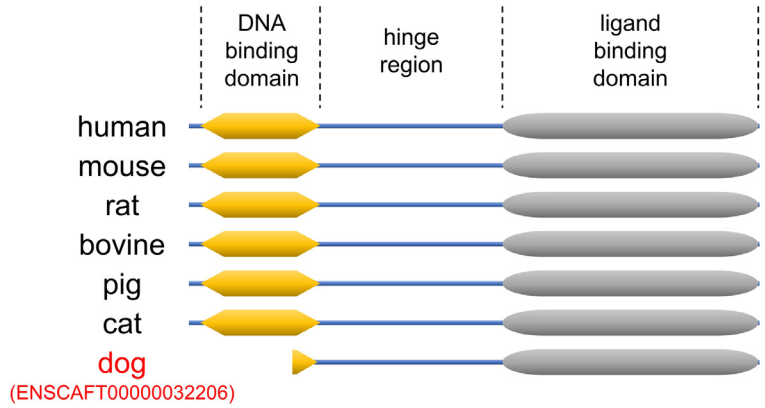

Fig. 1. Schematic overviews of Steroidogenic factor 1 (SF-1) registered in public databases of several species. In most species, SF-1 contains two domains: a DNA binding domain (DBD) and ligand binding domain (LBD) that are linked via a hinge region. Dog SF-1 (ENSCAFT00000032206) lacks the N-terminal aa sequence containing the major part of the DBD.

difference or database defect is unknown. However, considering the marked similarity of SF-1 among animals, except for dogs, and the fact that ENSCAFT00000032206 is computationally predicted, it is possible that the true exon1 and 2 of dog *SF-1* gene are veiled.

In humans and mice, the basal promoter of *SF-1* has been reported to contain a gene regulatory element with lower CpG density that exhibits dynamic DNA methylation patterns in cells or tissues [12]. DNA methylation generally occurs at the cytosine of CpG dinucleotides in higher vertebrates and is involved in tissue-specific and developmentally-regulated gene expression [25]. Expression of tissue-specific genes, such as *Oct-4* [11] and *Sry* [19], are induced by transient demethylation during development. The human and mouse *SF-1* gene is also considered to be under the control of DNA methylation. However, the regulation mechanism of dog *SF-1* gene is still unknown. Identification of the true promoter region of the dog *SF-1* gene will enable the analysis of the epigenetic regulation of the dog *SF-1* gene.

In this study, we identified the full-length dog *SF-1* cDNA and promoter sequence. We also examined the relationship between the expression levels of dog *SF-1* and the DNA methylation status in the basal promoter.

MATERIALS AND METHODS

Animals and sample collection

An adrenal gland from a male mixed breed dog, an ovary from a female beagle dog, a testis from a male beagle dog, and adipose tissues from a male chihuahua dog were collected during surgery at the University of Miyazaki Veterinary Teaching Hospital, Miyazaki, Japan, with the signed informed consent from dog owners and the ethical approval of the animal ethics committee of Faculty of Agriculture, University of Miyazaki, and the university's research committee. All samples were grossly normal. The tissues used in this study were as follows; a part of the adrenal gland including capsule, cortex, and medulla; a part of the ovary including germinal epithelium, tunica albuginea, cortex, and medulla; and a part of the testis including tunica albuginea, vascular layer, parenchyma, and interstitium. The adrenal gland, ovary, and testis were immediately frozen in liquid nitrogen and then stored at −80°C until use.

Isolation and culture of dog adipose tissue-derived mesenchymal stem cells (AD-MSCs)

Intra-abdominal adipose tissues were aseptically collected from a 9-year-old male chihuahua dog. Adipose tissues were cut into pieces ≤0.2 mm^3 and digested at 37°C with 0.1% (w/v) Trypsin-EDTA (FUJIFILM Wako Pure Chemical Corp., Tokyo, Japan). After digestion, the cell suspension was centrifuged at 1,000 rpm for 3 min to collect the cells. The cells were cultured in Dulbecco's modified Eagle's medium-low glucose (DMEM-LG; Sigma-Aldrich, St. Louis, MO, USA) supplemented with 10% (v/v) fetal bovine serum (Thermo Fisher Scientific, Waltham, MA, USA), 2 mM GlutaMAX™ Supplement (Thermo Fisher Scientific), 100 U/m*l* Penicillin-Streptomycin (Thermo Fisher Scientific), and 0.1% 2-mercaptoethanol (Thermo Fisher Scientific). After 48 hr, the medium was replaced, and non-adherent cells were removed. The medium was changed every 3 days. When the growth of the AD-MSCs was 70–80% confluent, the cells were detached by incubation in 0.05% trypsin-EDTA for 5 min at 37°C. For demethylation assays, AD-MSCs were cultured for 96 hr in medium containing 0, 1, 5, or 10 μM 5-aza-2′-deoxycytidine (5-aza-dC; Merck Millipore, Billerica, MA, USA).

Cloning and sequencing of full-length cDNA of dog SF-1

Total RNA was extracted from tissues and cells using ISOGEN II (FUJIFILM Wako Pure Chemical Corp.) following

the manufacturer's instructions. Quality and concentration were measured using a NanoDrop® 2000C spectrophotometer (Thermo Fisher Scientific). Rapid amplification of cDNA ends (RACE) reactions were performed using the GeneRacer™ Kit with SuperScript™ III RT (Thermo Fisher Scientific) and template cDNA from total RNA obtained from adrenal gland. The GeneRacer™ Kit provides a method to the obtain full-length 5′ and 3′ ends of cDNA by removing the mRNA cap structure, ligating the GeneRacer™ RNA Oligo to the mRNA, and reverse transcribing the mRNA with oligo dT primer. Specific amplification products were obtained through polymerase chain reaction (PCR) performed under the following thermocycling conditions: 30 cycles of 98°C for 10 sec, 60°C for 5 sec, and 72°C for 3 min. The primers used are summarized in Table 1. All PCR products were extracted with Wizard® SV Gel and PCR Clean-Up System (Promega, Madison, WI, USA) and were ligated into pBluescript II SK (−) by In-Fusion (TaKaRa Bio Inc., Kusatsu, Japan). Ligated PCR products were sub-cloned and sequenced.

Genomic DNA extraction and genome walking

Genomic DNA was isolated from tissues and cells by phenol and chloroform separation, and ethanol precipitation. This DNA was suspended in TE buffer (nacalai tesque, Kyoto, Japan). Quality and concentration were measured using the aforementioned NanoDrop® 2000C spectrophotometer. Genomic DNA isolated from adrenal grand was amplified using Splinkerette PCR [26] with specific primers and sequenced to determine the genomic locations of dog *SF-1*. All Splinkerette PCR experiments were performed under the following thermocycling conditions: 30 cycles of 98°C for 10 sec, 60°C for 5 sec, and 72°C for 3 min. The primers used are summarized in Table 1. All PCR products were sub-cloned and sequenced described above.

Bioinformatics analysis

Ensemble IDs of transcriptions used in this study were as follows: Homo sapiens (ENST00000373588.8), *Mus musculus* (ENSMUST00000112883.7), *Rattus norvegicus* (ENSRNOT00000017651.3), *Bos taurus* (ENSBTAT00000011869.3), *Sus scrofa* (ENSSSCT00000034748.2), *Felis catus* (ENSFCAT00000026158.3), and *Canis lupus familiaris* (ENSCAFT00000032206.3). Multiple alignments of SF-1 protein sequences were analyzed using clustalW (https://clustalw.ddbj.nig.ac.jp). The open reading frame (ORF) of dog *SF-1* was identified with ORF finder (https://www.ncbi.nlm.nih.gov/orffinder/). The amino acid sequence deduced from the cDNA sequence was obtained with the EMBOSS Transeq (https://www.ebi.ac.uk/Tools/st/emboss_transeq). Calculated molecular weights and predicted isoelectric points were obtained with EMBOSS Pepstats (https://www.ebi.ac.uk/Tools/seqstats/emboss_pepstats). Sequence identity of cDNA and basal promoter was analyzed using EMBOSS Stretcher (https://www.ebi.ac.uk/Tools/psa/emboss_stretcher) and EMBOSS Water (https://www.ebi.ac.uk/Tools/psa/emboss_water), respectively. TFBIND (http://tfbind.hgc.jp) was used to search for transcription factor binding sites in the dog *SF-1* basal promoter.

Table 1. Primers used in this study

Primers	Sequence (5′-3′)	Application
5′_dSF1-F	CGACTGGAGCACGAGGACACTGA	RACE
5′_dSF1-R	GTCCACGATGGAGATGAAGG	RACE
5′_Nested_dSF1-F	GGACACTGACATGGACTGAAGGAGTA	RACE
5′_Nested_dSF1-R	GCTCTGGGTACTCAGACTTGATG	RACE
3′_dSF1-F	TCCAGAAGTGCCTGACAGTG	RACE
3′_dSF1-R	GCTGTCAACGATACGCTACGTAACG	RACE
3′_Nested_dSF1-F	AGCATCTGGGCAACGAGATG	RACE
3′_Nested_dSF1-R	CGCTACGTAACGGCATGACAGTG	RACE
CDS_dSF1-F	ATGGACTATTCGTACGACGAG	RACE
CDS_dSF1-R	TCAAGTCTGCTTGGCTTGCA	RACE
Sp_adaptor	CGAAGAGTAACCGTTGCTAGGAGAGACC	Splinkerette PCR
Sp_Nested_adaptor	GTGGCTGAATGAGACTGGTGTCGAC	Splinkerette PCR
Sp_dSF1-F	CATGGACTATTCGTACGACGAGGACCTG	Splinkerette PCR
Sp_Nested_dSF1-F	GCTACCACTACGGACTGCTCACG	Splinkerette PCR
Sp_dSF1-R	ACCTTGCAGCTCTCGCACGTG	Splinkerette PCR
Sp_Nested_dSF1-R	ACGTGAGCAGTCCGTAGTGGTAGC	Splinkerette PCR
M13-F	TGTAAAACGACGGCCAGT	Sequence
M13-R	CAGGAAACAGCTATGACCATG	Sequence
Seq_dSF1-F	ACCGCACGCGCTGATATAG	Sequence
Seq_dSF1-R	ACGACAAAACCCCGATTCTGAG	Sequence
GAPDH-F	AATGCCTCCTGCACCACCAAC	qPCR
GAPDH-R	GAAGGCCATGCCAGTGAGCTTC	qPCR
dSF1-F	TCCAGAAGTGCCTGACAGTG	qPCR
dSF1-R	TGAAGCCATTGGCTCGAATCTG	qPCR
Bis_dSF1-F	GATTTAAATGAAGAGAAATATTAATAAAGAAGG	Bisulfite PCR
Bis_dSF1-R	ACCATAAACACATTCACAAACTAC	Bisulfite PCR

Gene expression analysis

Total RNA was extracted from tissues and cells using ISOGEN II (FUJIFILM Wako Pure Chemical Corp.) following the manufacturer's instructions. For the reverse transcription-polymerase chain reaction (RT-PCR), first-strand cDNA was synthesized using total RNA (1 μg) with random hexamers and ReverTra Ace reverse transcriptase (TOYOBO Co., Ltd., Osaka, Japan). The cDNA template was amplified using BIOTAQ^{TM} HS DNA Polymerase (Bioline Ltd.; London, UK) and specific primers for dog *SF-1* and dog *glyceraldehyde-3-phosphate dehydrogenase* (*GAPDH*). All PCR experiments were performed under the following thermocycling conditions: 95°C for 10 min; 30 cycles of 95°C for 30 sec, 60°C for 30 sec, and 72°C for 1 min, with a final extension at 72°C for 10 min. Quantitative real-time PCR (qPCR) was performed using SYBR® Green PCR master mix (Applied Biosystems, Woburn, MA, USA). Data were normalized to *GAPDH* expression. Gene expression levels are presented as the fold-change in expression, which was calculated using the Pfaffl method [22]. The sequences of the primers used in this study are summarized in Table 1.

Combined bisulfite restriction analysis (COBRA) and bisulfite sequencing

Sodium bisulfite treatment of genomic DNA was performed using the EZ DNA Methylation-Gold^{TM} Kit (Zymo Research, Irvine, CA, USA). PCR amplification was performed using BIOTAQ^{TM} HS DNA Polymerase (Bioline Ltd.) and specific primers for dog *SF-1* promoter. The sequences of primers used in this study are summarized in Table 1. All PCR experiments were performed under the following thermocycling conditions: 95°C for 10 min; 35 cycles of 95°C for 30 sec, 55°C for 30 sec, and 72°C for 1 min, with a final extension at 72°C for 10 min. For COBRA [32], the PCR product was treated by HpyCH4IV (New England Biolabs Inc., Ipswich, MA, USA) or TaqI (New England Biolabs Inc.). Concentration of the treated PCR products was measured using MultiNA (Shimadzu, Kyoto, Japan). To determine the methylation states of individual CpG sites at the dog *SF-1* promoter, the PCR product was gel-extracted, sub-cloned into the pGEM-T Easy vector (Promega), and sequenced. Methylation sites were visualized and quality control was performed using the QUMA web-based tool (http://quma.cdb.riken.jp/) [14].

Statistical analysis

Differences between two independent samples were evaluated by performing two-tailed Student's *t*-test. All error bars represent the standard error of the mean. Linear regression and Pearson product-moment correlation coefficient were used to analyze correlations between gene expression and DNA methylation.

RESULTS

Identification and sequence analysis of full-length dog SF-1 gene

To identify the true full-length sequence of dog *SF-1* mRNA, we extracted total RNA from dog adrenal grand and performed 5′- and 3′-RACE (Fig. 2a). In 5′-RACE PCR, five PCR products were obtained and sequenced. One of the five PCR fragments contained the dog *SF-1* sequence. The other fragments were non-specific products. In 3′-RACE PCR, four PCR products were obtained and sequenced. One of the four PCR fragments contained dog *SF-1* sequence. The other fragments were non-specific products. Combined the results of 5′- and 3′-RACE, a true full-length dog *SF-1* cDNA sequence was determined. No other mRNA variant was detected. The full-length dog *SF-1* cDNA was consisted of 3,016 base pairs (bp) including a 162 bp 5′ untranslated region (UTR), 1,386 bp ORF, and 1,468 bp 3′ UTR (Fig. 2b). The sequence was submitted to the DNA Data Bank of Japan (DDBJ) and assigned an accession number (ID: LC494495).

Since the full-length sequence of dog *SF-1* cDNA was clarified, we attempted to identify the dog *SF-1* position in the *Canis lupus familiaris* genome database (CanFam 3.1) using BLAT. However, most of the dog *SF-1* sequence identified in this study was not present in the database, indicating that the genomic sequence immediately upstream of the transcription start site (TSS) of ENSCAFT00000032206 is not registered. Therefore, we performed genome walking using Splinkerette PCR. Sequencing analysis of Splinkerette PCR products determined that the dog *SF-1* is located on chromosome 9 (estimated position: 58,460,020-58,484,999) and contains seven exons (Fig. 2c). The identified genomic sequence was submitted to DDBJ and assigned an accession number (ID: LC494496).

We compared the nucleic acid sequence between the full-length dog *SF-1* determined in this study and ENSCAFT00000032206 (Fig. 2c). The full-length dog *SF-1* cDNA was 300 bp long upstream of the predicted start codon of ENSCAFT00000032206. The 3′ UTR of the full-length dog *SF-1* cDNA was 1,716 bp, which was shorter than that of ENSCAFT00000032206. Other parts of the full-length dog *SF-1* cDNA sequence almost matched to those of ENSCAFT00000032206, except for some single nucleotide polymorphisms.

Based on the cDNA sequence, the dog SF-1 protein is composed of 461 amino acids, with an estimated molecular weight of 51.6 kDa and theoretical isoelectric point of 7.66. When the amino acid sequence of the full-length dog SF-1 was aligned to the SF-1 of several mammals, the sequence in DBD completely matched, demonstrating that SF-1 is also highly conserved in dogs (Fig. 3). The protein sequence of the full-length dog SF-1 shared 95.0 and 94.8% identity to the sequence of human and mouse, respectively. These data indicated that the full-length cDNA identified in this study is probably an original and normal sequence of dog *SF-1*.

Analysis of dog SF-1 basal promoter

Clarifying the TSS of dog *SF-1* identified the genome position and sequence of the basal promoter. The DNA sequence of the dog *SF-1* basal promoter (from −120 bp to +120 bp) shared 88.9% and 81.8% identity to that of human and mouse, respectively.

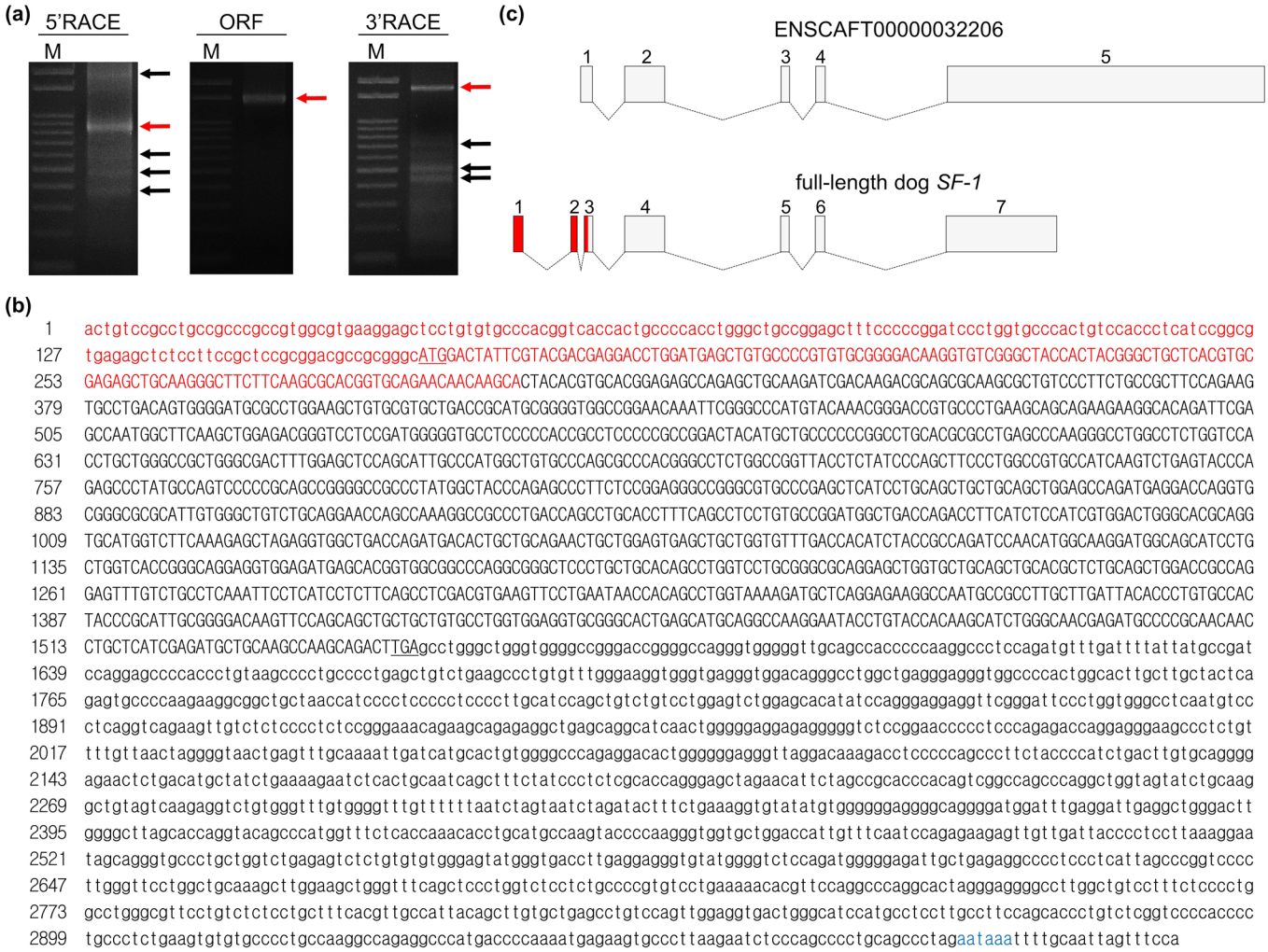

Fig. 2. (a) Representative electrophoresis image. All polymerase chain reaction (PCR) fragments (arrows) were sub-cloned and sequenced. Red arrows indicate the fragments including dog *Steroidogenic factor 1* (*SF-1*) mRNA sequence. Black arrows represent non-specific products. M: Marker. (b) The full-length cDNA sequence of dog *SF-1* identified in this study. Numbers on the left represent nucleotide positions. The 5′ and 3′ untranslated regions (UTRs) are indicated in lower case. The open reading frame is indicated in upper case. The start codon and the stop codon are underlined. Sequences in red represent the region newly identified in this study. Sequences in blue represent the polyadenylation signal. (c) Comparison of exon-intron structure of ENSCAFT00000032206 and the full-length dog *SF-1*. Exons are shown as boxes and are numbered. The coding region newly identified in this study are filled with red.

Similar to human and mouse *SF-1*, the basal promoter of dog *SF-1* contains regulatory elements including the SOX9 binding site, E box, CCAAT box, and Sp1/Sp3 site (Fig. 4a). DNA methylation at the basal promoter in the *SF-1* gene has been reported to regulate tissue-specific expression in humans and mice. Dog *SF-1* also has 15 CpG sites around exon 1 (Fig. 4a). Thus, dog *SF-1* may be also under the control of DNA methylation.

Relationship between DNA methylation at dog SF-1 promoter and gene expression

To investigate whether DNA methylation is involved in the regulation of dog *SF*-1 expression, we performed demethylation assay with the demethylating reagent 5-aza-dC using dog AD-MSCs. Dog *SF-1* mRNA was not detected in the AD-MSCs. However, 5-aza-dC treatment induced the dose-dependent expression of dog *SF-1* (Fig. 4b and 4c). This result suggested that DNA methylation affects the expression of dog *SF-1*.

We next analyzed the expression of dog *SF-1* gene and DNA methylation around exon 1 in steroidogenic tissues, including adrenal, ovary, and testis. The gene expression was detected in adrenal, ovary, and testis, but the expression level was tissue-dependent (Fig. 5a and 5b). The DNA methylation rates of three CpG sites around the dog *SF-1* promoter, −82 bp, + 286 bp, and + 426 bp from TSS, was analyzed using COBRA. DNA methylation rates were 0–3% in ovary, 30–39% in adrenal, 63–69% in testis, and 74–100% in AD-MSCs (Fig. 5c). There was a clear inverse correlation between the gene expression and the DNA methylation rates among samples (Fig. 5d). Bisulfite sequencing analysis was performed to investigate the DNA methylation state of individual CpG sites around the dog *SF-1* promoter from positions −93 bp to + 426 bp, which contains 30 CpG sites (Fig. 5e). As expected,

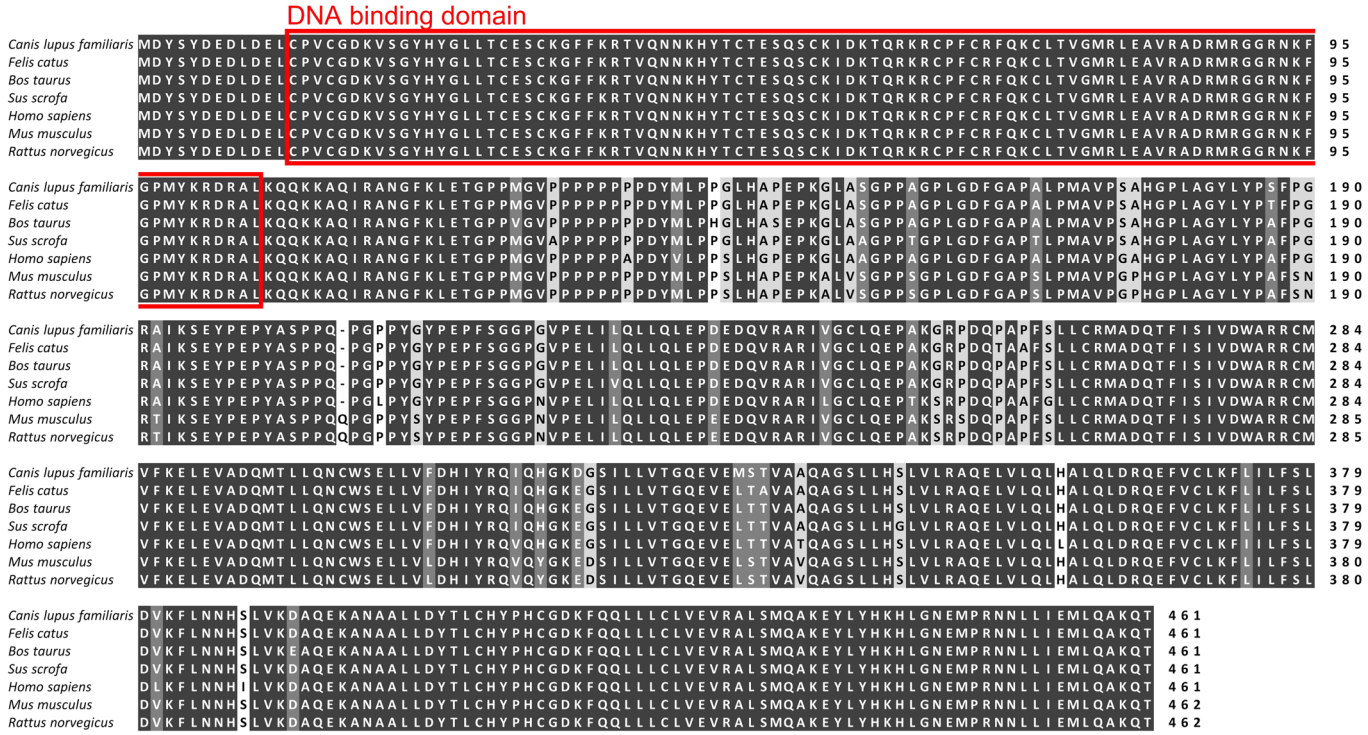

Fig. 3. Multiple sequence alignments of deduced amino acid (aa) sequence of full-length dog Steroidogenic factor 1 (SF-1) with other species. Alignments were performed with clustalW. Conserved aa sequences are indicated by a dark background. Highly similar aa sequences are indicated by a dark grey background. Weakly similar aa sequences are indicated by a light grey background. Numbers on the right show the position of the aa sequence. DNA binding domain is boxed in red.

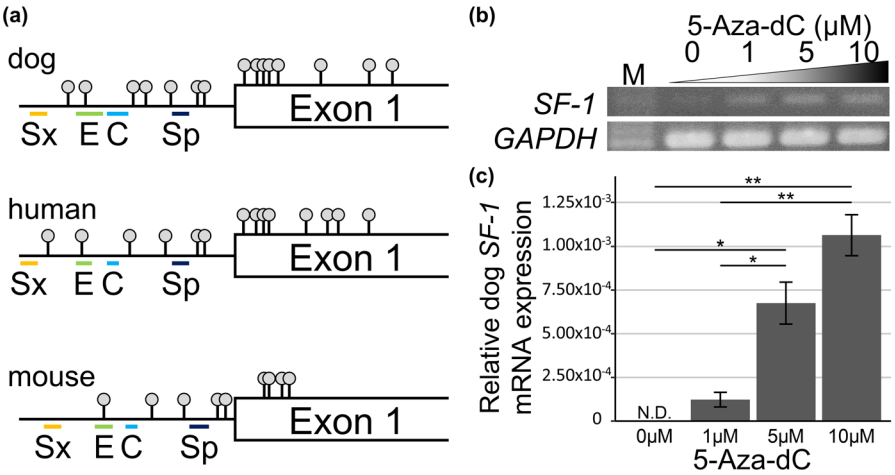

Fig. 4. (a) Overview of the CpG sites in the *Steroidogenic factor 1* (*SF-1*) basal promoter (−120 pb ~ +120 bp) in dogs, humans, and mice. Lollipops indicate the position of individual CpG sites. Horizontal lines indicate transcription factor binding sites. Sx, a binding site for SRY-box9 (SOX9); E, an E box; C, a CCAAT box; Sp, a binding site for Sp1 or Sp3 transcription factors. (b, c) Effect of a demethylating reagent (5-aza-dC) on the expression of dog *SF-1* gene in adipose tissue-derived mesenchymal stem cells (AD-MSCs). (b) Expression of the dog *SF-1* and dog *glyceraldehyde-3-phosphate dehydrogenase* (*GAPDH*) genes by reverse transcription-polymerase chain reaction (RT-PCR). Expression of the dog *SF-1* increased depending on 5-aza-dC concentration. *GAPDH* was used as an internal control. (c) Relative *SF-1* gene expression levels measured by quantitative RT-PCR, with normalization to *GAPDH* expression using the Pfaffl method. The data shown represent the mean ± standard error (SE) (n=3). **P*<0.05, ***P*<0.01. N.D.: not detected.

based on the COBRA results, methylation levels were low throughout the promoter in ovary, while AD-MSCs were highly methylated. In addition, the inverse correlation of methylation pattern to expression extended downstream of exon 1. These results indicated that the promoter activity of dog *SF-1* is under the control of DNA methylation.

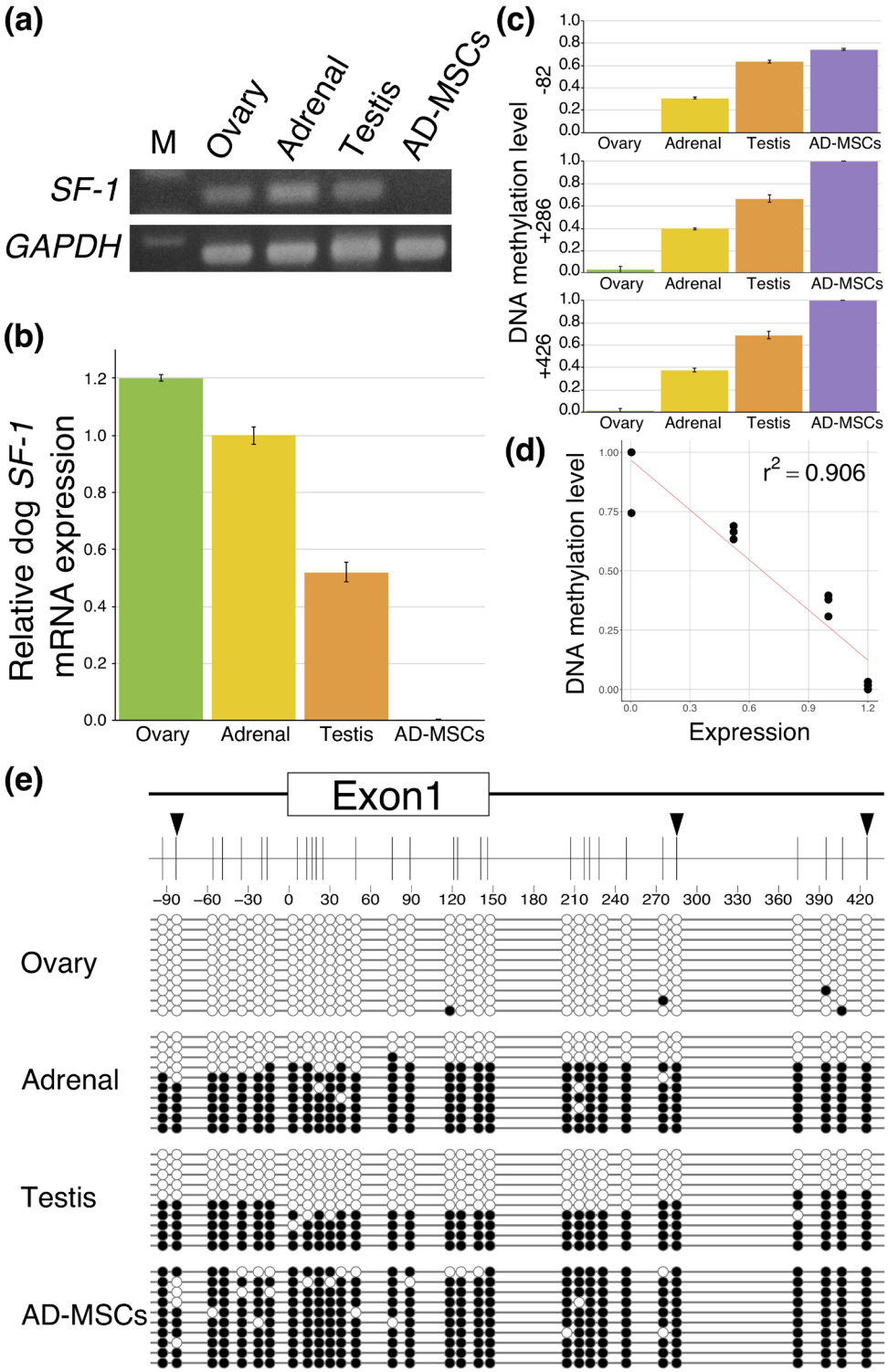

Fig. 5. Analysis of dog *Steroidogenic factor 1* (*SF-1*) expression and DNA methylation in the promoter. (a) Expression of the dog *SF-1* gene in adrenal gland, ovary, testis and adipose tissue-derived mesenchymal stem cells (AD-MSCs) by reverse transcription-polymerase chain reaction (RT-PCR). The upper and lower panels show dog *SF-1* and dog *glyceraldehyde-3-phosphate dehydrogenase* (*GAPDH*) expression, respectively. (b) Relative dog *SF-1* gene expression levels measured by quantitative RT-PCR with normalization to *GAPDH* expression using the Pfaffl method. The data shown represent the mean ± SE (n=3). (c) DNA methylation level measured by combined bisulfite restriction analysis (COBRA) at the dog *SF-1* promoter region in tissues and cultured cells. Numbers on the left indicate the position of CpG site from TSS. The data shown represent the mean ± SE (n=3). (d) Scatter plot of dog *SF-1* gene expression and DNA methylation levels in the promoter, defined by quantitative RT-PCR and COBRA. Red line indicates linear regression line. (e) Bisulfite sequencing analysis around exon 1 of dog *SF-1* gene. (Top) Diagram of the dog *SF-1* gene. Exon 1 is shown as a white box. Vertical lines indicate the position of individual CpG sites. Black arrowheads represent the position of the CpG sites measured in COBRA (d). (Bottom) The open and closed circles indicate the unmethylated and methylated states of each CpG site, respectively.

DISCUSSION

The dog *SF-1* mRNA recorded in the database (CanFam3.1) lacks the 5′-end sequence coding for the DBD, which is conserved in other mammals. In this study, we identified the true full-length cDNA of dog *SF-1*. It possessed the 5′-end sequence coding DBD and shared high similarities with sequences in human, mouse, bovine, pig, and cat. In addition, determination of the 5′-end sequence of dog *SF-1* mRNA enabled us to identify the genomic location and the genomic sequence of the promoter. There are many unclarified genomic sequences in the dog genome database [13]. Most of the RefSeq genes in dog have been computationally predicted based on CanFam3.1. In fact, 3,181 RefSeq genes in CanFam3.1 have unclarified regions within 5,000 upstream from TSS, suggesting that the correct 5′-end sequences of those genes have been veiled.

Many previous studies have reported that point mutations of the human *SF-1* gene cause adrenal insufficiency and disorders of sex development (DSD). In particular, mutations in DBD, such as p.G35E or p.R92Q, led to severe phenotype of those diseases [8]. The 46, XY DSD, which is the most common *SF-1*-related disease featuring a DBD mutation, includes clitoral enlargement, small inguinal testes, and absent or rudimentary Müllerian structures as the typical phenotype. In the field of veterinary medicine, dog 78, XY DSD exhibits symptoms including testicular hypoplasia with clitoral enlargement, persistent Müllerian duct syndrome, cryptorchidism, and hypospadias, similar to human 46, XY DSD [17, 23]. However, dog *SF-1* gene mutations have not been detected in the dog 78, XY DSD. One of the reasons is that the genome sequence of dog *SF-1* DBD have not been identified. Our results provide the sequence of dog *SF-1* DBD and information that will be useful in veterinary medicine diagnosis and research.

In this study, the basal promoter sequence of dog *SF-1* gene was identified. Epigenetic analyses of the dog *SF-1* promoter revealed that the expression of dog *SF-1* gene is under the control of DNA methylation. These results indicate that epigenetic mutation influences gene expression of *SF-1* in dog. Aberrant hypomethylation at the basal promoter of the *SF-1* gene has been reported to induce ectopic gene expression in human endometriosis [33]. In dogs, ectopic-endometrium and endometrioma have been reported, which are homologous diseases to human endometriosis [1, 2, 7, 21]. However, the relationship between those dog diseases and *SF-1* has not been clarified. Based on our findings, it is possible that ectopic *SF-1* gene is overexpressed in dog endometriosis by epigenetic mutation in the promoter.

In conclusion, the complete sequences of mRNA and the promoter region of dog *SF-1* were identified. Expression of dog *SF-1* is under the control of DNA methylation at the promoter. Our results provide a molecular biological basis for a better understanding of developmental and metabolic mechanisms for dogs.

CONFLICTS OF INTEREST. The authors declare that they have no conflict of interest.

ACKNOWLEDGMENTS. We would like to express our sincere thanks to Mr. Takuro Naora, Mr. Shotaro Mori for generously providing us with materials. This research was supported by a grant-in-aid from the University of Miyazaki to KN.

REFERENCES

1. Bartel, C., Berghold, P. and Walter, I. 2011. Ectopic endometrial tissue in mesonephric duct remnants in bitches. *Reprod. Domest. Anim.* **46**: 950–956. [Medline] [CrossRef]
2. Bartel, C., Schönkypl, S. and Walter, I. 2010. Pseudo-placentational endometrial cysts in a bitch. *Anat. Histol. Embryol.* **39**: 74–80. [Medline] [CrossRef]
3. Bassett, M. H., Suzuki, T., Sasano, H., White, P. C. and Rainey, W. E. 2004. The orphan nuclear receptors NURR1 and NGFIB regulate adrenal aldosterone production. *Mol. Endocrinol.* **18**: 279–290. [Medline] [CrossRef]
4. Bassett, M. H., Zhang, Y., Clyne, C., White, P. C. and Rainey, W. E. 2002. Differential regulation of aldosterone synthase and 11beta-hydroxylase transcription by steroidogenic factor-1. *J. Mol. Endocrinol.* **28**: 125–135. [Medline] [CrossRef]
5. Clyne, C. D., Zhang, Y., Slutsker, L., Mathis, J. M., White, P. C. and Rainey, W. E. 1997. Angiotensin II and potassium regulate human CYP11B2 transcription through common cis-elements. *Mol. Endocrinol.* **11**: 638–649. [Medline] [CrossRef]
6. Daggett, M. A., Rice, D. A. and Heckert, L. L. 2000. Expression of steroidogenic factor 1 in the testis requires an E box and CCAAT box in its promoter proximal region. *Biol. Reprod.* **62**: 670–679. [Medline] [CrossRef]
7. Demirel, M. A. 2017. A case of spontaneous abortion related to ovarian endometriosis in a Golden Retriever dog. *Majallah-i Tahqiqat-i Dampizishki-i Iran* **18**: 63–66. [Medline]
8. El-Khairi, R. and Achermann, J. C. 2012. Steroidogenic factor-1 and human disease. *Semin. Reprod. Med.* **30**: 374–381. [Medline] [CrossRef]
9. Hammer, G. D., Krylova, I., Zhang, Y., Darimont, B. D., Simpson, K., Weigel, N. L. and Ingraham, H. A. 1999. Phosphorylation of the nuclear receptor SF-1 modulates cofactor recruitment: integration of hormone signaling in reproduction and stress. *Mol. Cell* **3**: 521–526. [Medline] [CrossRef]
10. Hanley, N. A., Rainey, W. E., Wilson, D. I., Ball, S. G. and Parker, K. L. 2001. Expression profiles of SF-1, DAX1, and CYP17 in the human fetal adrenal gland: potential interactions in gene regulation. *Mol. Endocrinol.* **15**: 57–68. [Medline] [CrossRef]
11. Hattori, N., Nishino, K., Ko, Y. G., Hattori, N., Ohgane, J., Tanaka, S. and Shiota, K. 2004. Epigenetic control of mouse Oct-4 gene expression in embryonic stem cells and trophoblast stem cells. *J. Biol. Chem.* **279**: 17063–17069. [Medline] [CrossRef]
12. Hoivik, E. A., Aumo, L., Aesoy, R., Lillefosse, H., Lewis, A. E., Perrett, R. M., Stallings, N. R., Hanley, N. A. and Bakke, M. 2008. Deoxyribonucleic acid methylation controls cell type-specific expression of steroidogenic factor 1. *Endocrinology* **149**: 5599–5609. [Medline] [CrossRef]
13. Holden, L. A., Arumilli, M., Hytönen, M. K., Hundi, S., Salojärvi, J., Brown, K. H. and Lohi, H. 2018. Assembly and analysis of unmapped genome sequence reads reveal novel sequence and variation in dogs. *Sci. Rep.* **8**: 10862. [Medline] [CrossRef]
14. Kumaki, Y., Oda, M. and Okano, M. 2008. QUMA: quantification tool for methylation analysis. *Nucleic Acids Res.* **36**: W170-5. [Medline] [CrossRef]

15. Luo, X., Ikeda, Y. and Parker, K. L. 1994. A cell-specific nuclear receptor is essential for adrenal and gonadal development and sexual differentiation. *Cell* **77**: 481–490. [Medline] [CrossRef]
16. Martin, L. J. and Tremblay, J. J. 2005. The human 3beta-hydroxysteroid dehydrogenase/Delta5-Delta4 isomerase type 2 promoter is a novel target for the immediate early orphan nuclear receptor Nur77 in steroidogenic cells. *Endocrinology* **146**: 861–869. [Medline] [CrossRef]
17. Meyers-Wallen, V. N. 2012. Gonadal and sex differentiation abnormalities of dogs and cats. *Sex Dev.* **6**: 46–60. [Medline] [CrossRef]
18. Morohashi, K., Honda, S., Inomata, Y., Handa, H. and Omura, T. 1992. A common trans-acting factor, Ad4-binding protein, to the promoters of steroidogenic P-450s. *J. Biol. Chem.* **267**: 17913–17919. [Medline]
19. Nishino, K., Hattori, N., Tanaka, S. and Shiota, K. 2004. DNA methylation-mediated control of Sry gene expression in mouse gonadal development. *J. Biol. Chem.* **279**: 22306–22313. [Medline] [CrossRef]
20. Nomura, M., Bärtsch, S., Nawata, H., Omura, T. and Morohashi, K. 1995. An E box element is required for the expression of the ad4bp gene, a mammalian homologue of ftz-f1 gene, which is essential for adrenal and gonadal development. *J. Biol. Chem.* **270**: 7453–7461. [Medline] [CrossRef]
21. Paiva, B. H., Silva, J. F., Ocarino, N. M., Oliveira, C. A., Assis, W. A. and Serakides, R. 2015. A rare case of endometrioma in a bitch. *Acta Vet. Scand.* **57**: 31. [Medline] [CrossRef]
22. Pfaffl, M. W. 2001. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* **29** No. 9e: e45. [Medline] [CrossRef]
23. Poth, T., Breuer, W., Walter, B., Hecht, W. and Hermanns, W. 2010. Disorders of sex development in the dog-Adoption of a new nomenclature and reclassification of reported cases. *Anim. Reprod. Sci.* **121**: 197–207. [Medline] [CrossRef]
24. Ramayya, M. S., Zhou, J., Kino, T., Segars, J. H., Bondy, C. A. and Chrousos, G. P. 1997. Steroidogenic factor 1 messenger ribonucleic acid expression in steroidogenic and nonsteroidogenic human tissues: Northern blot and in situ hybridization studies. *J. Clin. Endocrinol. Metab.* **82**: 1799–1806. [Medline] [CrossRef]
25. Razin, A., Webb, C., Szyf, M., Yisraeli, J., Rosenthal, A., Naveh-Many, T., Sciaky-Gallili, N. and Cedar, H. 1984. Variations in DNA methylation during mouse cell differentiation in vivo and in vitro. *Proc. Natl. Acad. Sci. USA* **81**: 2275–2279. [Medline] [CrossRef]
26. Shao, H. and Lok, J. B. 2014. Detection of piggyBac-mediated transposition by splinkerette PCR in transgenic lines of Strongyloides ratti. *Bio Protoc.* **4**: e1015. [Medline] [CrossRef]
27. Shen, J. H. and Ingraham, H. A. 2002. Regulation of the orphan nuclear receptor steroidogenic factor 1 by Sox proteins. *Mol. Endocrinol.* **16**: 529–540. [Medline] [CrossRef]
28. Sugawara, T., Kiriakidou, M., McAllister, J. M., Kallen, C. B. and Strauss, J. F. 3rd. 1997. Multiple steroidogenic factor 1 binding elements in the human steroidogenic acute regulatory protein gene 5′-flanking region are required for maximal promoter activity and cyclic AMP responsiveness. *Biochemistry* **36**: 7249–7255. [Medline] [CrossRef]
29. Ueda, H. and Hirose, S. 1991. Defining the sequence recognized with BmFTZ-F1, a sequence specific DNA binding factor in the silkworm, Bombyx mori, as revealed by direct sequencing of bound oligonucleotides and gel mobility shift competition analysis. *Nucleic Acids Res.* **19**: 3689–3693. [Medline] [CrossRef]
30. Wang, X. L., Bassett, M., Zhang, Y., Yin, S., Clyne, C., White, P. C. and Rainey, W. E. 2000. Transcriptional regulation of human 11beta-hydroxylase (hCYP11B1). *Endocrinology* **141**: 3587–3594. [Medline] [CrossRef]
31. Woodson, K. G., Crawford, P. A., Sadovsky, Y. and Milbrandt, J. 1997. Characterization of the promoter of SF-1, an orphan nuclear receptor required for adrenal and gonadal development. *Mol. Endocrinol.* **11**: 117–126. [Medline] [CrossRef]
32. Xiong, Z. and Laird, P. W. 1997. COBRA: a sensitive and quantitative DNA methylation assay. *Nucleic Acids Res.* **25**: 2532–2534. [Medline] [CrossRef]
33. Yamagata, Y., Nishino, K., Takaki, E., Sato, S., Maekawa, R., Nakai, A. and Sugino, N. 2014. Genome-wide DNA methylation profiling in cultured eutopic and ectopic endometrial stromal cells. *PLoS One* **9**: e83612. [Medline] [CrossRef]
34. Zeitoun, K., Takayama, K., Michael, M. D. and Bulun, S. E. 1999. Stimulation of aromatase P450 promoter (II) activity in endometriosis and its inhibition in endometrium are regulated by competitive binding of steroidogenic factor-1 and chicken ovalbumin upstream promoter transcription factor to the same cis-acting element. *Mol. Endocrinol.* **13**: 239–253. [Medline] [CrossRef]